Available Online at ESci Journals
Journal of Plant Breeding and Genetics
ISSN: 2305-297X (Online), 2308-121X (Print)
http://www.escijournals.net/JPBG

MATING DESIGNS: HELPFUL TOOL FOR QUANTITATIVE PLANT BREEDING ANALYSIS

[a,b]Athanase Nduwumuremyi*, [a]Pangirayi Tongoona, [c]Slyvestre Habimana

[a] *African Centre for Crop Improvement (ACCI), University of KwaZulu-Natal, Scottville, Pietermaritzburg, South Africa.*

[b] *Department of Natural Resource Management, Rwanda Agriculture Board (RAB), Kigali, Rwanda.*

[c] *Faculty of Agriculture and rural development, Higher Institute of Agriculture and Animal Husbandry, Musanze, Rwanda.*

A B S T R A C T

Selection of parental materials and good mating designs in conventional plant breeding are the keys to the successful plant breeding programme. However, there are several factors affecting the choices of mating designs. Mating design refers to the procedure of producing the progenies, in plant breeding, plant breeders and geneticists, theoretically and practically, they use different form of mating designs and arrangements for targeted purpose. The choice of a mating design for estimating genetic variances should be dictated by the objectives of the study, time, space, cost and other biological limitations. In all mating designs, the individuals are taken randomly and crossed to produce progenies which are related to each other as half-sibs or full-sibs. A form of multivariate analysis or the analysis of variance can be adopted to estimate the components of variances. Therefore, this review aimed at highlighting the most used mating design in plant breeding and genetics studies. It provides easy and quick insight of the different form of mating designs and some statistical components for successful plant breeding.

Keywords: mating designs, plant breeding, genetics statistics.

INTRODUCTION

In plant breeding, various mating designs and arrangements are used by breeders and geneticists to generate improved plants. The selection of suitable parents and good mating designs are keys to the successful plant breeding schemes (Khan *et al*., 2009). However, there are several factors affecting the choices of mating designs. The factors influencing the choice of mating design are (i) the type of pollination (self- or cross-pollinated); (ii) the type of crossing to be used (artificial or natural); (iii) the type of pollen dissemination (wind or insect); (iv) the presence of a male-sterility system; (v) the purpose of the project (for breeding or genetic studies); and (vi) the size of the population required (Acquaah, 2012).

Before discussing the mating designs, it is very important to understand the genetic assumptions (Hill *et al*., 1998) : (a) Diploid behaviour at meiosis; this assumption applies to all designs, but it doesn't preclude the investigation of polyploidy species provided they behave as functional diploids, with disomic inheritance. (b) Uncorrelated genes distribution. The genes controlling the character should be independently distributed among the parents. (c) Absence of non-allelic interactions. In the triple test and diallel crosses epistasis can be detected and its effects including in prediction. (d) No multiple alleles at those loci controlling the character. (e) Absence of reciprocal differences. Again this assumption can be tested in several designs and appropriate measure taken. (f) Ideally the diallel cross should be restricted to crosses among homozygous lines. Heterozygous can be catered for, but it complicated the interpretation of the results. (g) Absence of genotype-environment interaction. Their presence merely emphasizes the need for wide scale testing of material in order to determine the extent of such interaction.

The mating designs have four main importance, (1) to provide information on the genetic control of the character under investigation; (2) to generate a breeding population to be used as a basis for the selection and

* Corresponding Author:
Email ID: nduwatha@gmail.com

development of potential varieties; (3) to provide estimates of genetic gain and (4) to provide information for evaluating the parents used in the breeding program (Acquaah, 2012). In making various crosses breeders have interests in discovering the answer to the following questions: how significant is genetic variation? How much of the variation is heritable? And what types of gene affecting that significance? However, these are answered by comparing the variances of the segregating and the non-segregating generations (Kearsey and Pooni, 1996). Another interest of the breeder is identifying plants with superior genotypes as judged by the performance of their progeny. Suitable inbreds or lines are selected based on combining ability effects with better mean performance. Combining ability depends on the gene action controlling the trait to be improved. General combining ability (GCA) is the average performance of a line in hybrid combinations and is due to additive genes action. The estimation of GCA for a particular line depends upon the mating design, but essentially, it is the deviation of its progeny mean from the mean of all lines included in the trial (Acquaah, 2012). Thus, theoretically, differences between maternal groups measure variation in their general combining ability. Specific combining ability (SCA) refers to combinations or crosses that do relatively better or worse than would be expected based on the average performance of the lines involved, it is therefore due to non-additive gene action (Falconer and Mackay, 1996, Acquaah, 2012). The information regarding the estimates of combining ability and genes actions is vital for a successful plant breeding (Panhwar *et al*., 2008). Plant breeding experiments use two types of design, (1) mating and (2) experimental design which should march with its statistical components analysis and interpretation. Therefore, this review provides the different form of mating designs and some statistical components for successful plant breeding; these will serve as reference to the scientists and students in the domain of plant breeding and genetics.

MAJOR MATING DESIGNS IN PLANT BREEDING AND GENETICS

Mating design refers to the procedure of producing the progenies, in plant breeding, plant breeders and geneticists, theoretically and practically, they use different form of mating designs and arrangements for targeted purpose. However, the choice of a mating design for estimating genetic variances should be dictated by the objectives of the study, time, space, cost and other biological limitations. Thus, several studies (Griffing, 1956b; Kearsey and Pooni, 1996; Hallauer *et al*., 2010; Acquaah, 2012) described and contrasted different mating designs and six types of mating designs have been described so far: (1) bi-parental progenies (BIP), polycross, topcross, North Carolina (I, III, III), Diallels (I, II, III, IV) and Line X tester design. In all mating designs, the individuals are taken randomly and crossed to produce progenies which are related to each other as half-sibs or full-sibs. A form of multivariate analysis or the analysis of variance can be adopted to estimate the components of variances.

BI-PARENTAL MATING

The bi-parental design is also called paired crossing design and it is reported to be the simplest mating design (Mather, 1982). In this design, the breeder selects a large number of plants (n) at random and cross them in pairs to produce 1/2n full-sib families (Acquaah, 2012). Their progeny are tested and the observed variation partitioned by straightforward analysis of variance into between and within families (Hill *et al*., 1998). Statistically, if r plants per progenies family are evaluated, the variation within (w) and between (b) families may be analyzed as follows:

Table 1. Analysis of variance for BIP design.

Source of variation	df	MS	EMS
Between families	$(\frac{a}{b}n)-1$	MS_1	$\sigma^2 w + r\sigma^2 b$
Within families	$\frac{a}{b}n(r-1)$	MS_2	$\sigma^2 w$
Total	$\frac{a}{b}nr-1$		

Source: Acquaah, 2012

Where: n and r refer to the number of parents and plants sampled within each cross respectively;

$\sigma^2 b$ is the covariance of full-sibs ($\sigma^2 b$ = CovFS = $\frac{1}{2}V_A + \frac{1}{4}V_D + V_{EC} = \frac{1}{r}(MS_1 - MS_2)$ and

$\sigma^2 w = \left[\sigma^2_G - CovFS\right] + \sigma^2_{EW} = \frac{1}{2}V_A + \frac{3}{4}V_D + V_{EW} = MS_2$; is the environmental source of variation for variance within the crosses. When you assume that dominance effects are zero, then $\sigma^2 b = \frac{1}{2}V_A$ and $\sigma^2 w = \frac{1}{2}V_A + V_{EW}$ (Acquaah, 2012).

The simplicity of this design is counterbalanced by its inability to yield sufficient information to estimate all parameters required by the model (Acquaah, 2012). Only two statistics are available for estimating V_A, V_D, V_{EW} and V_{EC}. This is because the progeny from this design are either full sibs or unrelated; no other relationship exists among them. The estimates of the parameters can only be obtained either by simplifying assumptions, or if extra statistics become available (Hill *et al*., 1998). If dominance is assumed to be absent (V_D=0), and there is no common environment (V_{EW}=0), that is individuals from the same family do not share the same environment. Consequently, if these assumptions are unjustified, it will lead to an overestimate of the genetic component relative to the environmental component. These difficulties can be circumvented to a limited extent in practice. Family plots can be dispensed by randomizing individual plants over the whole experiment. In this way V_{EC} becomes zero. For biometrical geneticist individual plant randomization is a useful device for increasing the precision for estimates, but for a plant breeder it may be an unaffordable luxury. Nevertheless, for shrubs and trees, single tree plot designs are frequently used. Usually, however, it may be simpler for the breeder to estimate V_{EC} directly from the families X replicates interaction mean square in properly replicated experiment.

In BIP design, extra statistic can be generated if information on the parents is available, or by inclusion of the selfed progenies of parents (Kearsey, 1970). However, in practice both options are of limited value; because of the problems posed by presence of genotypes-environment interaction when considering parent-offspring correlation, and introduction of additional parameters.

The BIP design, though simple to execute, has obvious limitations. Because no constraints have been imposed upon the mating there is a lack of relatedness among the resultant progeny. Consequently, unjustifiable assumptions many be required if estimated of the most important genetic and environmental components are to be obtained (Hill *et al*., 1998). The most limitation of this design is its inability to provide the needed information to estimate all the parameters required by the model. The progeny from the design comprise full sibs or unrelated individuals. There is no further relatedness among individuals in the progeny. The breeder must make unjustifiable assumptions in order to estimate the genetic and environmental variance.

POLYCROSS

This design is for intermating a group of cultivars by natural crossing in isolated block. Term polycross was coined by Tysdal, Kiesselbach and Westover in 1942, to indicate progeny from seed of a line that was subject to outcrossing with other selected lines growing in the same nursery (Hill *et al*., 1998). It is most suited to species that are obligate cross-pollinaters (e.g., forage grasses and legumes, sugarcane, sweet potato), but especially to those that can be vegetatively propagated crops such as sugarcane, cassava and sweet potato (Acquaah, 2012). The design provides equal opportunity for each and every clone or parent to naturally cross with each other in the block such that self-pollination is prevented (Saladaga, 1989). However, to achieve this objective, a proper design in the polycross block is critical. It provides an equal opportunity for each entry to be crossed with every other entry. It is critical that the entries be equally represented and randomly arranged in the crossing block (Falconer and Mackay, 1996). The Latin square experimental design was suggested to be used as the most appropriate design to ensure all entries have equal chance of random intermating with each other in the polycross nursery (Morgan, 1988). Nevertheless, when the entries number is more than 10, the completely randomized block design may be used (Acquaah, 2012). In both cases, about 20–30 replications are included in the crossing block. The ideal requirements are hard to meet in practice because of several problems, placing the system in jeopardy of deviating from random mating. If all the entries do not flower together, mating will not be random. To avoid this, the breeder may plant late flowering entries earlier. Pollen may not be dispersed randomly, resulting in concentrations of common pollen in the crossing block. Half sibs are generated in a polycross because progeny from each entry has a common parent. The design is used in breeding to produce synthetic cultivars, recombining selected entries of families in recurrent selection breeding programs, or for evaluating the GCA of entries (Acquaah, 2012).

The polycross design has an advantage of producing synthetic cultivars, recombining selected genotypes in the recurrent selection procedure and evaluating the general combing ability of the parent genotypes (Falconer and Mackay, 1996; Sleper and Poehlman, 2006; Acquaah, 2012). The general combining ability

(GCA) can also be estimated from polycross mating design.

The general combining abilities estimated are basically for maternal parents and the variations measured in a progeny can be partitioned into within and between maternal parents (Falconer and Mackay, 1996) and consequently, general combining ability helps in estimating heritability. The mean performance of the progenies of any female parent in the polycross is used to determine the variance components and consequently the general combining ability (GCA). The heritability calculated provides decision guidance for usefulness of polycross in breeding programme (Saladaga, 1989). However, since the parents are of different origin and the crop is sensitive to environmental changes, the performance of the parental lines and their progenies such as flowering is likely to be affected (Morgan, 1988). In addition, the differences in performance of progeny clones could arise from variations in heritability of trait measured (Gorz and Haskins, 1971). Consequently these could lead to inaccurate estimates of GCA; hence heritability determined needs to be treated with caution.

In polycross, the progenies from individual plants are tested. These progenies are half-sib families. The covariance within the families is $Cov(HS) = \frac{1+F}{4}\sigma^2{}_A$ where F the inbreeding coefficient of the genotypes being tested. The following ANOVA table shows the analysis of many replications for polycross design.

Table 1. ANOVA table of polycross design with many replicated.

Source	df	MS	EMS	Variance components
Progenies	$g-1$	M_1	$\sigma_e^2 + r\sigma_{prog}^2$	$\sigma^2 prog = Cov(HS) = \frac{1+F}{4}\sigma^2 A$
Blocks	$r-1$	M_2	-	-
error	$(g-1)(r-1)$	M_3	σ_e^2	$\sigma^2{}_e = \sigma^2$

Source: Wricke and Weber, 1986

The variance component s_{prog}^2 is an estimate of $\frac{1+F}{4}\sigma_A^2$; when the parents are non-inbred, F = zero. A comparison of the coefficients with the corresponding coefficients in case of parent-offspring covariance indicates that the precision of the estimate of σ_A^2 is lower for the topcross or polycross than for the covariance between parents and offspring. The precision is increased if the tested genotypes are inbred. It is convenient to use polycross design in cross-pollinated species when evaluating a large number of genotypes (Wricke and Weber, 1986). The selection is then applied based on half-sib progeny means. However, polycross design has a number of limitation such as random mating, insufficient statistics to estimate all the parameters, the component of variance are only estimated from the maternal half sibs; information about the males is lost, no control over the pollen source; expected genetic gains are reduced by half, the non-randomness of mating (due to lack of synchronisation of flowering, unequal pollen production and position effects in the crossing block). The polycross is ideally suited for identifying mother plants with superior genotypes from the performance of their progeny general combining ability (Hill *et al*., 1998).

TOP CROSS DESIGN

Topcross refers to a mating between a selection, line, clone and a common pollen parent which may be a variety, inbred line or single cross. The selected plants are crossed with a common tester(s) of known performance, generally in open pollination. The design was proposed by Jenkins and Brunsen in 1932 for testing inbred lines of maize in cross-bred combinations and later renamed topcross by Tysdal and Grandall in 1948 (Hill *et al*., 1998). The tester parent should have well known genetic background; either narrow- or broad-based testers (Aly *et al*., 2011). The purpose of using top is to increase the chance of obtaining a desirable gene or genes from exotic or difficult materials. Exotic refers to lines from other countries which are generally poorly adapted to local conditions. Difficult material refers to varieties or lines which are tall, poor combiners, or dominant susceptibles, etc. i.e. lines which have given poor results (progeny) from single crosses in previous crossing cycles. In making top crosses, only single cross F1's are utilized because they are uniform. The top cross F1's will be segregating and it is impossible to identify superior plants at crossing; therefore, they are not used. The F1's are selected for

desirable agronomic characteristics or for desirable parentage.

Topcross has been fairly widely used for preliminary evaluation of combining ability of new inbred lines (Mosa, 2010). The possible numbers of crosses are *n* x 1, given *n* number of inbreds. Topcross progenies yield only GCA information, not SCA. It is also called inbred-variety cross (Sleper and Poehlman, 2006). It is a simple and efficient system of screening inbred lines for combining ability before pairing inbreds in single-cross yield trials. This design is probably the simplest model of mating design that can provide preliminary rapid screening of genetic stocks as it involves the lowest crossing load and simple statistical analysis (Mosa, 2010). In topcross, the progenies from individual plants are tested; these progenies are half-sib families. The covariance within the families is $Cov(HS) = \frac{1+F}{4}\sigma_A^2$ where $F$the inbreeding coefficient of the genotypes is tested. Table 3 presents the general ANOVA for topcross.

Table 3. Skeleton of ANOVA for half sib family test by topcross.

Source	df	MS	EMS	Variance components
Progenies	$g-1$	M_1	$\sigma_e^2 + r\sigma_{prog}^2$	$\sigma^2 prog = Cov(HS) = \frac{1+F}{4}\sigma^2 A$
Blocks	$r-1$	M_2	-	
Error	$(g-1)(r-1)$	M_3	σ_e^2	$\sigma^2{}_e = \sigma^2$

Source: Wrickle and Weber, 1986.

The variance component $\sigma^2{}_{prog}$ is an estimate of $\frac{1+F}{4}\sigma^2 A$, calculated from $\sigma^2{}_{prog} = V(m_1) + V(m_2)$, when the parents are non-inbred, F = zero (Wrickle and Weber, 1986). However, top crosses require 5 heads per cross; this number is necessary because these crosses will segregate in the next F1 generation and at least 80 plants to facilitate the selection of desirable plants in the F_1. The design has two shortfalls. First, a single tester variety may not offer wide genetic background for testing the inbred stocks. Secondly, the numbers of crosses become large if the test inbreds are many.

NORTH CAROLINA

North Carolina design was developed after using long time diallel. However, the later require much labour. Therefore, in order to obtain more information about combining ability but without much labour comparing to full diallel, Comstock and Robinson in 1952, introduced the North Carolina designs I, II, and III.

North Carolina Design I: It is a very popular multipurpose design for both theoretical and practical plant breeding applications (Acquaah, 2012). It is commonly used to estimate additive and dominance variances as well as for the evaluation of full- and half-sib recurrent selection. It requires sufficient seed for replicated evaluation trials, and hence is not of practical application in breeding species that are not capable of producing large amounts of seed. It is applicable to both self- and cross-pollinated species that meet this criterion. As a nested design (Figure 1), each member of a group of parents used as males is mated to a different group of parents. NC Design I is a hierarchical design with non-common parents nested in common parents (Acquaah, 2012).

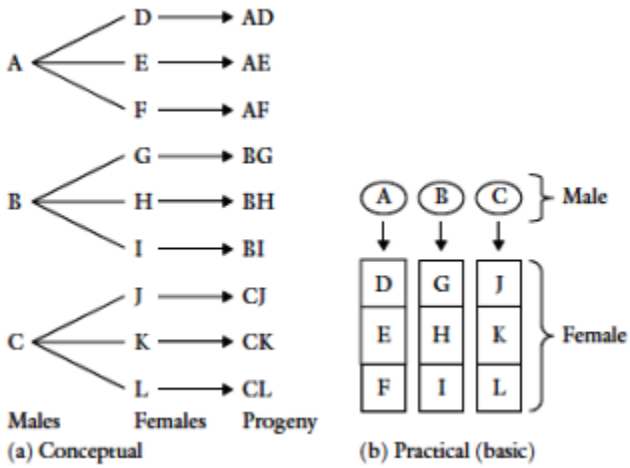

Figure 1. North Carolina Design I (Source: Acquaah, 2012).

The progenies include both full-sibs and half-sibs. Each set of families with the same father in common constitutes a half sib family group and a set of families with both parents in common constitutes a full-sib family (Kearsey and Pooni, 1996; Hallauer *et al*., 2010). The model for experiment conducted in one environment is: $Y_{ijk} = \mu + m_i + f_{ij} + r_k + e_{ijk}$ Where

μ is the mean, m_i is the effect of the *i*th male, f_{ij} is the effect of the *j*th female mated to the *i*th male, r_k is the replication effects, and e_{ijk} is the experimental error

(Hallauer *et al*., 2010). The design is commonly used to estimate additive and dominance variances (Acquaah, 2012).

Table 2. Skeleton of general ANOVA for North Carolina design I.

Source of variation	df	MS	Expected mean squares
Males	$(m-1)$	M_1	$\sigma^2 w + r\sigma^2_{mf} + rf\sigma^2_m$
Females	$m(f-1)$	M_2	$\sigma^2 w + r\sigma^2_{mf}$
Within plots	$mf(r-1)$	M_1	$\sigma^2 w$
Total	$rmf-1$		

Source: Acquaah, 2012.

The parameter $\sigma^2 w$ refers to the average variance within the full sib families and is given as;

$$\sigma^2 w = M_1 = \tfrac{1}{2}V_A + \tfrac{3}{4}V_D + V_E$$

$$\sigma^2_m = (M_1 - M_2)/rf = \tfrac{1}{4}V_A$$

$$r\sigma^2_{mf} = (M_2 - M_3)/r = \tfrac{1}{4}V_A + \tfrac{1}{4}V_D$$

The translation of components of variance to covariance of relatives permits estimation of components of genetic variances (σ^2_A and σ^2_D) and their interaction with the environments. Assuming no epistasis, the estimation of additive genetic variance (σ^2_A) and total genetic variance (σ^2_G) are obtained from the mean squares of the analysis of variance (Hallauer *et al*., 2010). Dominance variance σ^2_D is obtained from the difference between the female - within males and males components of variance. The NCI has advantage over biparental and polycross designs, because it gives three statistics compared with only two in the polycross and biparental (Kearsey, 1965).

Like the polycross, the main advantage of design I is its ability to supply a test of significance for the additive genetic variance. Also, NCI has been used successfully to tree breeding where mass collection of pollen from common parents possess no practical problems (Hill *et al*., 1998). Further, the design is applicable to both self- and cross-pollinated crops. However, this design is most widely used in animal studies. In plants, it has been extensively used in maize breeding for estimating genetic variances (Acquaah, 2012).

North Carolina Design II : In this design, each member of a group of parents used as males is mated to each member of another group of parents used as females Design II is a factorial mating scheme (Figure 2). It is used to evaluate inbred lines for combining ability.

The design is most adapted to plants that have multiple flowers so that each plant can be used repeatedly as both male and female.

Blocking is used in this design to allow all mating involving a single group of males to a single group of females to be kept intact as a unit (Acquaah, 2012). The design is essentially a two-way ANOVA in which the variation may be partitioned into difference between males (m) and females (f) and their interaction.

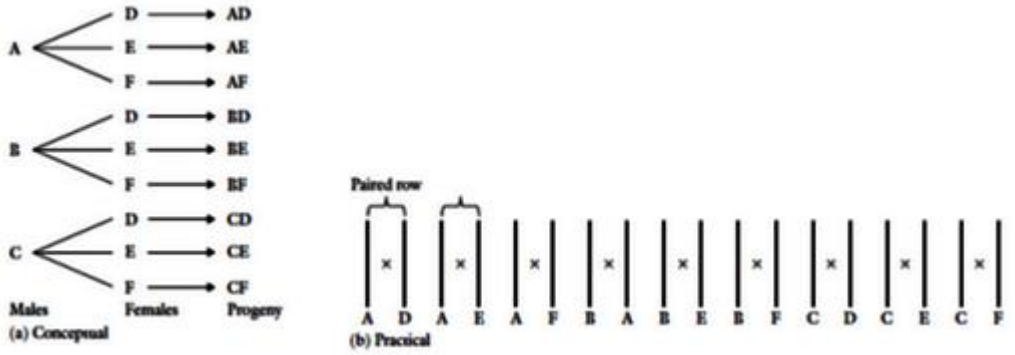

Figure 2: NC II Design (factorial design with paired rows) Source: Acquaah, 2012.

Table 5. Skeleton of general ANOVA of NC II.

Source of variation	df	MS	Expected mean squares
Replications	$r-1$		
Males	$m-1$	M_1	$\sigma^2 w + r\sigma^2_{mf} + rf\sigma^2_m$
Females	$f-1$	M_2	$\sigma^2 w + r\sigma^2_{mf} + rm\sigma^2_f$
Males x females	$(m-1)(f-1)$	M_3	$\sigma^2 w + r\sigma^2_{mf}$
Within progenies	$mf(r-1)$	M_4	$\sigma^2 w$
Error	$(r-1)(mf-1)$	M_5	σ^2
Total	$rmf-1$		

Source: Kearsey and Pooni, 1996

Where: $\sigma^2 w$ is the within progenies genetic and environmental variances. In the absence of epistasis and common environmental effects, σ^2_{mf} is a function of dominance variance V_D only (Kearsey and Pooni, 1996). If there is environmental variation between FS families, this could be due to general and specific maternal effects. The general maternal effects (V_{EM}) will appear in $\sigma^2_f = \frac{1}{4}V_A + V_{EM}$, while the specific maternal effects $V_{EC} - V_{EM}$ will appear in $\sigma^2_{mf} = \frac{1}{4}V_D + (V_{EC} - V_{EM})$ and will be confounded with V_D (Kearsey and Pooni, 1996). If the number of males and females is the same, $n_1 = n_2 = n$, we can have a test of maternal effects by comparing M_1/M_2 as a variance ratio.

This design also allows the breeder to measure not only GCA but also SCA (Acquaah, 2012). However, the NCII is not providing test of epistasis or G X E interaction (Kearsey and Pooni, 1996). In North Carolina II, every progeny family has half sib relationships through both common male and common female. This is accomplished by systematic crossing programme in which n_1 male and n_2 female are mated in all possible combinations to give n_1n_2 progeny families. It is therefore a rectangular mating design, unless n_1=n_2. Reciprocal crosses may be carried out to analyze maternal effects (Hill, *et al*., 1998).

North Carolina Design III: In this design, a random sample of F2 plants is backcrossed to the two inbred lines from which the F2 was descended. It is considered the most powerful of all the three NC designs.

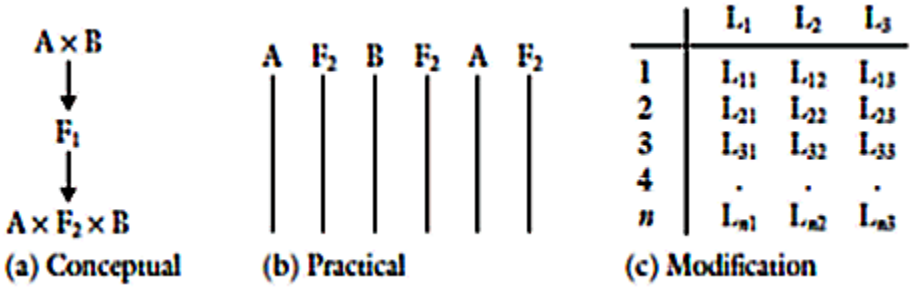

Figure 3. NC III design with conceptual, practical and modified. Source: Acquaah, 2012

However, it was made more powerful by the modifications made by Kearsey and Jinks that add a third tester not just the two inbreds (Acquaah, 2012). The two parental lines act as testers against which F_2 are assessed. The parents being progenitors of the F_2, are very special testers because F_2 is segregating at all loci for which the testers differ but for no other loci (Kearsey and Pooni, 1996). The F2 population is reference population for mating NCIII (Hallauer *et al*., 2010).

The modification is called the triple testcross and is capable of testing non-allelic (epistatic) interactions, which the other designs cannot, and also capable of estimating additive and dominance variance (Acquaah, 2012). According to Hill *et al*, 1998, it is also called triple test cross because of inclusion of the third tester. This inclusion increase the power of this design considerably, because it provides a sensitive and unambiguous test for non-allelic interactions, a capability which none of the designs described so far, not even design 3 in its original form possess. Moreover, both in its original and

extended form design 3 has a general utility for investigating any population, irrespective of gene frequency or mating system (Hill *et al*., 1998). In triple test cross a random sample of n individuals from the population under investigation is crossed to the same three testers, L1, L2 and L3 to give 3n progeny families. The analysis of this design may be divided into two parts, the first part supplies a test for epistasis, and the second assesses the significance, and provides estimates of the additive and dominance components of variation. The NCIII is a special case of NCII, therefore the ANOVA is similar to that of the NCII although it differs in one special feature; the two testers are not a random sample from any population but are two very particular lines, the progenitor of the F_2.

Table 7. Skeleton of NC III ANOVA.

Source of variation	df	MS	Expected mean squares
Testers, p	1	M_4	$\sigma^2 + r\sigma^2_{mp} + rmK^2_p$
Males (F_2), m	$m-1$	M_3	$\sigma^2 + 2r\sigma^2_m$
Testers x parents	$m-1$	M_2	$\sigma^2 + r\sigma^2_{mp}$
Within FS families/error	$(r-1)(2m-1)$	M_1	σ^2
Total	$2mr-1$		

Source: Hallauer *et al*., 2010

Where; r and m are number of replications and male plants, respectively. Although this design had many advantages, there are some pitfalls which should be avoided if its full potential is to be realized. If interest centres on determining whether dominance is present in particular population, virtually any pair of inbred lines will suffice testers. But if we wish to estimate the total dominance variation, L1 and L2 should differ at all those loci which are segregating in the population. The same criterion also applies if an unbiased estimate of the additive component is required. Generally speaking therefore, L1 and L2 should be high and low selections from the population, when additive and dominance components will be estimated with equal precision. But the tester must be derived from the population under investigation as the model developed for this mating design is only valid within a population. Because L1 and L2 represent extreme selections, estimates of narrow sense heritability can be obtained only after selection had taken place (Hill *et al*., 1998). It is reported that the main use of the triple test cross will be to investigate the inheritance of quantitative traits in natural population and relating this to natural selection and ecology (Kearsely and Jinks, 1968).In practice the triple test cross is viewed as a resource consuming design. Moreover, the generations required to establish this design are not those which a breeder would routinely produce in a breeding programme. For the biometrical geneticist, interest in determining the genetic architecture of a particular character and in estimating the genetic component with maximum precision, the triple test cross is without doubt the most powerful of the available mating designs. But for the plant breeder this design may be off-piste (Hill *et al*., 1998).

DIALLEL DESIGN

A complete diallel mating design is one that allows the parents to be crossed in all possible combinations (Schlegel, 2010), including selfs and reciprocals. This is the kind of mating scheme required to achieve Hardy–Weinberg equilibrium in a population (Acquaah, 2012).

The diallel is the most used and abused of all mating designs in obtaining various genetic information (Hallauer *et al*., 2010). Much of its abuse could probably be due to the presence of two models for diallel analysis; random and fixed models (Griffing, 1956b). A random model involves parents that are random members of a random mating population. A random model is useful for estimating GCA and SCA variances. In contrast, when parents are considered fixed effects, the aim is to measure the GCA effect for each parent and the SCA effect for each pair of parents. These effects only apply to the set of parents in the diallel. It is also widely used for developing breeding populations for recurrent selection (Acquaah, 2012). In addition, Johnson and King (1998), reported that diallel mating designs are deployed to provide the maximum opportunity to manage co-ancestry in breeding population and maximize selection differential. However, in practice, a diallel with selfs and reciprocals is neither practical nor useful for several reasons. Selfing does not contribute to the recombination of genes between parents. Furthermore, recombination is achieved by crossing in one direction

making reciprocals unnecessary (Acquaah, 2012). Because of the extensive mating patterns, the number of parents that can be mated this way is limited.

Nursery arrangements for the application vary depending either complete or partial diallel design and four methods under the diallel mating design have been so far described. The number of progenies generated from each method are different, the number of progeny families (pf) for methods 1 through 4 are: pf = n2, pf = 1/2n (n + 1), pf = n(n − 1), and pf = 1/2n(n − 1), respectively (Acquaah, 2012).

Method I or full diallel design: The method I or full diallel design consisted by parents, one set of F_1's and reciprocal F_1's. The system gives n^2 genotypes (Griffing, 1956b). The mathematical models for combining ability analysis for the fixed and random effects are given by;

Fixed effect model or model I:

$$Y_{ij} = \mu + g_i + g_j + s_{ij} + r_{ij} + \frac{1}{bc}\sum_k (bv)_{ijk} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl}$$

The table 8 presents the estimates of variances for variance components both fixed and random model.

Table 8. Skeleton of ANOVA for method I diallel design.

Source	df	SS	MS	Expected mean squares	
				Model I	Model II
GCA	$p-1$	S_g	M_g	$\sigma^2 + 2p(\frac{1}{p-1})\sum g^2{}_i$	$\sigma^2 + \frac{2(p-1)}{p}\sigma^2{}_g + 2p\sigma^2{}_g$
SCA	$p(p-1)/2$	Ss	M_s	$\sigma^2 + \frac{2}{p(p-1)}\sum\sum S_{ij}{}^2$	$\sigma^2 + \frac{2(p^2-p+1)}{p^2}\sigma_s{}^2$
Reciprocal eff.	$p(p-1)/2$	S_r	M_r	$\sigma^2 + 2(\frac{2}{p(p-1)})\sum_{i<j} \sum r_{ij}{}^2$	$\sigma^2 + 2\sigma^2{}_r$
Error	m	S_e	M_e	σ^2	

Source: Griffing (1956b)

The various effects can be estimated as follows:

$\hat{\mu} = \frac{1}{p^2}X..,\ \hat{g}_i = \frac{1}{2p}(Xi. + X.i) - \frac{1}{p^2}X..,$

$\hat{s}_{ij} = \frac{1}{2}(x_{ij} + x_{ji}) - \frac{1}{2p}(X_{i.} + X_{.i} + X_{j.} + X_{.j}) + \frac{1}{p^2}X..,$

$\hat{r}_{ij} = \frac{1}{2}(x_{ij} - x_{ji})$ (Griffing, 1956b). It is important to note these restrictions; $\sum_i gi = 0$ and $\sum_j s_{ij} + s_{ii} = 0$ for each i.

However, the variances of the effects can be estimated using the following equations: $Var(\hat{\mu}) = \frac{1}{p2}\hat{\sigma}^2,$

$Var(\hat{g}_i) = \frac{p-1}{2p^2}\hat{\sigma}^2,\ Var(\hat{s}_{ij}) = \frac{1}{2p^2}(p^2 - 2p + 2)\hat{\sigma}^2,\ i \neq j,$

$Y_{ij} = \mu + g_i + g_j + s_{ij} + r_{ij} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl}$ Where,

μ is the population mean, g_i, g_j is the general combining ability effect for the i^{th} and j^{th} parents, s_{ij} is the specific combining ability effect of the cross between the i^{th} and j^{th} parents such that $s_{ij} = s_{ji}$, r_{ij} is the reciprocal effect involving the reciprocal crosses between the i^{th} and j^{th} parents such that $r_{ij} = r_{ji}$ and, e_{ijkl} is the experimental error due to environmental effect associated with the $ijkl^{th}$, which is assumed to be uncorrelated and normally distributed with zero mean and variance, V_E

Random effect model or Model II:

and $Var(\hat{r}_{ij}) = \frac{1}{2}\hat{\sigma}^2$

Method II: This method includes parents and one set of F_1's without reciprocals F_1's. This design gives *p(p+1)/2* genotypes. The mathematical models for combining ability for fixed model is:

$$Y_{ij} = \mu + g_i + g_j + s_{ij} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl} \quad \begin{cases} i, j = 1,\ldots, p, \\ k = 1,\ldots, b, \\ l = 1,\ldots, c, \end{cases}$$

Where, μ is the population mean, g_i, g_j is the general combining ability effect for the i^{th} and j^{th} parents, s_{ij} is the specific combining ability effect of the cross between the i^{th} and j^{th} parents such that $s_{ij} = s_{ji}$ and

e_{ijkl} is the experimental error due to environmental effect associated with the *ijkl*th . It is important to remember that $\sum_i g_i = 0$ and $\sum_j s_{ij} + s_{ii} = 0$.

The mathematical equation for analysis of combining ability for random model is:

$$Yij = \mu + g_i + g_j + s_{ij} + \frac{1}{b}\sum_k b_k + \frac{1}{b}\sum_k (bv)_{ijk} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl},$$

Table 9. The analysis of variance for method II.

Source	df	SS	MS	Expected mean squares	
				Model I	Model II
GCA	$p-1$	S_g	M_g	$\sigma^2 + (2+p)(\frac{1}{p-1})\sum g_i^2$	$\sigma^2 + \sigma_s^2 + (p+2)\sigma_g^2$
SCA	$p(p-1)/2$	Ss	M_s	$\sigma^2 + \frac{2}{p(p-1)}\sum_i \sum_j s_{ij}^2$	$\sigma^2 + \sigma_s^2$
Error	m	S_e	M_e'	σ^2	

Source: Griffing (1956b)

The various effects for method II can be estimated as follows,

$$\hat{\mu} = \frac{2}{p(p+1)} X..,$$

$$\hat{g}_i = \frac{1}{p+2}\left[X_{i.} + x_{ii} - \frac{2}{p} X..\right],$$

$$\hat{s}_{ij} = x_{ij} - \frac{1}{p+2}\left[X_{i.} + x_{ii} + X_{j.} + x_{jj}\right] + \frac{2}{(p+1)(p+2)} X..$$

However, the variances of the effects can be estimated using the following equations: $Var(\hat{\mu}) = \frac{2}{p(p+1)}\hat{\sigma}^2$,

$$Var(\hat{g}_i) = \frac{p-1}{p(p+2)}\hat{\sigma}^2,$$

$$Var(s_{ii}) = \frac{p(p-1)}{(p+1)(p+2)}\hat{\sigma}^2 \text{ and } Var(s_{ij}) = \frac{p2+p+2}{(p+1)(p+2)}\hat{\sigma}^2$$

Method III: In this method, one set of F1's and the reciprocals are included. This mating design gives rise to $a = p(p-1)$ different number of genotypes. As for methods I and II, also it has both fixed and random effect models.

Fixed model:

$$Y_{ij} = \mu + g_i + g_j + s_{ij} + r_{ij} + \frac{1}{bc}\sum_k b_k + \frac{1}{b}\sum_k (bv)_{ijk} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl} + e_{ijkl},$$

$$\begin{cases} i, j = 1,\ldots, p, \\ k = 1,\ldots; b, \\ l = 1,\ldots, c, \end{cases}$$

μ is the population mean, g_i, g_j is the general combining ability effect for the *i*th and *j*th parents, s_{ij} is the specific combining ability effect of the cross between the *i*th and j th parents such that $s_{ij} = s_{ji}$ and r_{ij} is the reciprocal genotypic effects such that $r_{ij} = r_{ji}$ and, e_{ijkl} is the experimental error due to environmental effect associated with the *ijkl*th observation (Griffing, 1956b). However, the following equations help in estimating effects:

$$Var(\hat{\mu}) = \frac{1}{p(p-1)}\hat{\sigma}^2,\ Var(\hat{g}_i) = \frac{p-1}{2p(p-2)}\hat{\sigma}^2,$$

$$Var(\hat{s}_{ij}) = \frac{p-3}{2(p-1)}\hat{\sigma}^2 \ (i \neq j),\ Var(\hat{r}_{ij}) = \frac{1}{2}\hat{\sigma}^2 \ (i \neq j).$$

Random model:

$$Y_{ij} = \mu + g_i + g_j + s_{ij} + r_{ij} + \frac{1}{b}\sum_k b_k + \frac{1}{b}\sum_k (bv)_{ijk} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl},$$

Table 10. Skeleton of ANOVA of Diallel method III.

Source	df	SS	MS	Expected mean squares	
				Model I	Model II
GCA	$p-1$	S_g	M_g	$\sigma^2 + 2p(p-2)(\frac{1}{p-1})\sum g^2{}_i$	$\sigma^2 + 2\sigma_s^2 + 2(p-2)\sigma_g^2$
SCA	$p(p-3)/2$	Ss	M_s	$\sigma^2 + \frac{2}{p(p-3)}\sum_{i<j}\sum s_{ij}^2$	$\sigma^2 + 2\sigma_s^2$
Reciprocal eff.	$p(p-1)/2$	S_r	M_r	$\sigma^2 + 2(\frac{2}{p(p-1)})\sum_{i<j} \sum r_{ij}^2$	$\sigma^2 + 2\sigma_r^2$
Error	m	S_e	M_e'	σ^2	σ^2

Source: Griffing (1956b)

The variances for variance components can be estimated by the following equations,

$$Var(\hat{\sigma}_g^2) \cong \frac{1}{2p(p-1)(p-2)^2} M_g^2 + \frac{1}{p(p-2)^2(p-3)} M_s^2$$

$$Var(\hat{\sigma}_s^2) \cong \frac{1}{p(p-1)} M_s^2 + \frac{1}{2m}(M_e')^2,$$

$$Var(\hat{\sigma}_r^2) \cong \frac{1}{p(p-1)} M_r^2 + \frac{1}{2m}(M_e')^2,$$

$$Var(\hat{\sigma}^2) \cong \frac{2}{m}(M_e')^2,$$

2.5.4 Method IV

In this method, only one set of F1's are included. It is the most common of the diallel crossing systems. There are $a = p(p-1)/2$ different genotypes evaluated. As for other diallel methods, there are two models.

Fixed model:

$$Yij = \mu + g_i + g_j + s_{ij} + \frac{1}{bc}\sum_k \sum_l \varepsilon_{ijkl} \quad \begin{cases} i, j = 1,\ldots, p, \\ k = 1,\ldots, b \\ l = 1,\ldots, c \end{cases}$$, where

μ is the population mean, g_i, g_j is the general combining ability effect for the i^{th} and j^{th} parents s_{ij} is the specific combining ability effect of the cross between the i^{th} and jth parents such that $s_{ij} = s_{ji}$ and e_{ijkl} is the experimental error effect unique to the $ijkl^{th}$ observation (Griffing, 1956b). The variances of the effects may be estimated as follows:

$$Var(\hat{\mu}) = \frac{2}{p(p-1)}\hat{\sigma}^2, Var(\hat{g}_i) = \frac{p-1}{p(p-2)}\hat{\sigma}^2, \text{ and}$$

$$Var(\hat{s}_{ij}) = \frac{p-3}{p-1}\hat{\sigma}^2 \quad (i \neq j).$$

Table 11. Skeleton of ANOVA for Diallel method IV.

Source	df	SS	MS	Expected mean squares	
				Model I	Model II
GCA	$p-1$	S_g	M_g	$\sigma^2 + (p-2)(\frac{1}{p-1})\sum_i g_i^2$	$\sigma^2 + 2\sigma^2_s + (p-2)\sigma_g^2$
SCA	$p(p-3)/2$	Ss	M_s	$\sigma^2 + \frac{2}{p(p-3)}\sum_{i<j}\sum s_{ij}^2$	$\sigma^2 + \sigma_s^2$
Error	m	S_e	M_e'	σ^2	σ^2

Source: Griffing (1956b)

This mating design provides information on GCA and SCA (Griffing, 1956b). However, the fixed model of method 3 or 4 is the most appropriate for obtaining unbiased estimates of combining abilities and gene action (Shattuck *et al*., 1993). This method is most suitable when there are no genotypic reciprocal effects (Griffing, 1956b). The most of the problems arising with diallel crosses are essentially due to experimental design such that analysis of data is complex (Johnson and King, 1998).

LINE X TESTER DESIGN

Line x tester is basically an extension of topcross design in the sense that instead of one tester as used in topcross, more than ones testers are used under L x T mating design. Line x tester mating design was first proposed by Kempthorne in 1957 cited by Sharma (2006). This design involves hybridization between lines (f) and wide based testers in one to one fashion generating f x m = fm hybrids (Sharma, 2006). It is the simplest mating design that provides both full-sibs and half-sibs simultaneously as opposed to topcross which provides only half-sibs. It provides SCA of each cross, and it is not providing GCA of lines only but of the testers also, as liner and tester both are different sets of genotypes (Sharma, 2006). It is therefore most suitable for animal experiment (Sharma, 2006). In addition, it is used in estimating various types of gene actions important in the expression of quantitative traits (Rashid *et al*., 2007).

Its statistical model is: $Y_{ijkl} = \mu + a_l + b_{kl} + v_{ij} + (av)_{ijl} + \varepsilon_{ijkl}$, where Y_{ijkl} = observed value from each experimental unit; μ = population mean; a_l = location effect; b_{kl} = block or replication effect within each location; v_{ij} = F1 hybrid effect = $g_i + g_j + s_{ij}$. However, $v_{ij} = g_i + g_j + s_{ij}$, where g_i = general combining ability (GCA) for the i^{th} parental line; g_j = GCA effect of j^{th} tester;

s_{ij} = specific combining ability (SCA) for the ij^{th} F1 hybrid and $(av)_{ijl}$ = interaction effect between i^{th} F1 hybrid and l^{th} location; and ε_{ijkl} = residual effect.

Table 3. Skeleton of ANOVA for Line X Tester Design

Source	df	MS	Expected mean squares	
			Model I	Model II
Replication	$r-1$			
Lines	$m-1$	M_1	$\sigma^2 + rf\frac{1}{m-1} + \sum_i g_i^2$	$\sigma^2 + v_{sca} + rf_{gca(m)}$
Testers	$f-1$	M_2	$\sigma^2 + rm(\frac{1}{f-1})\sum_j g_j^2$	$\sigma^2 + rv_{sca} + rm_{gca(f)}$
Line x tester	$(m-1)(f-1)$	M_3	$\sigma^2 + r\left[\frac{1}{(m-1)(f-1)}\right]\sum_i\sum_j s_{ij}$	$\sigma^2 + rv_{sca}$
Error	$(r-1)(mf-1)$	M_4	σ^2	σ^2

Source: Sharma (2006)

The significance of mean square for line x testers provides a direct test of significance of dominance variance, σ^2_D while significance of σ^2_A is provided by significance of lines and testers mean squares.

CONCLUSION

Selection of suitable parents and good mating designs are keys to the successful of plant breeding schemes (Khan *et al*., 2009). However, selection of mating design depends on many factors, for instance in early stages of breeding programme, there are a large numbers of different lines, clones or accessions to be evaluated. In this case, appropriate designs would include the polycross and perhaps the topcross but the polycross is particularly suited to the identification of potential parents of synthetic varieties. A comparative evaluation of mating designs summarized mating designs in two ways (1) in terms of coverage of the population and it was reported that BIPs > NC I > polycross > NC III > NC II > diallel, in that order of decreasing effectiveness; (2) in terms of amount of information, diallel > NC II > NC III > NC I > BIPs (Acquaah, 2012). However, diallel mating design is the most important for GCA and SCA but breeder has a great role in choosing mating design instead of relying on the mating design *per se*. The proper choice and use of a mating design will provide the most valuable information for breeding (Acquaah, 2012).

REFERENCES

Acquaah, G. 2012. Principles of plant genetics and breeding. 2nd ed. Wiley-Blackwell, Oxford.

Aly, R.S.H., E.M.R. Metweli and S.T.M. Mousa. 2011. Combining ability of maize (*Zea mays* L.) inbred lines for grain yield and some agronomic traits using topcross mating design. Glob. J. Mol. Sci. 6: 1-8.

Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Pearson Prentice Hall, Harlow, England.

Gorz, H.J. and F.A. Haskins. 1971. Evaluation of cross-fertilization in forage crops. Agronomy--Faculty publications, University of Nebraska. Lincoln. p. 731-734.

Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci. 9: 463-493.

Hallauer, A.R., M.J. Carena and J.B.M. Filho. 2010. Quantitative genetics in maize breeding. 6th ed. Springer, Iowa, USA.

Hill, J., H.C. Becker and P.M.A. Tigerstedt. 1998. Quantitative and ecological aspects of plant breeding. Chapman and Hall, UK, London.

Johnson, G.R. and J.N. King. 1998. Analysis of half diallel mating designs: I - A practical analysis procedure for ANOVA approximation. Silvae Genetica. 47: 74-79.

Kearsely and Jinks. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits: I. Theory. Heridity. 23: 403-409.

Kearsey. 1970. Experiment sizes for detecting dominance variation. Heredity 24: 45-57.

Kearsey, M.J. 1965. Biometrical analysis of a random mating population: a comparison of five experimental designs. Heredity. 20: 205-235.

Kearsey, M.J. and H.S. Pooni. 1996. The genetical analysis of quantitative traits. 1st ed. Chapman and Hall, London, UK.

Khan, S.A., H. Ahmad, A. Khan, M. Saeed, S.M. Khan and B. Ahmad. 2009. Using line x tester analysis for earliness and plant height traits in sunflower

(*Helianthus annuus* L.). Recent Research in Science and Technology. 1: 202–206.

Mather, K.J., J. L 1982. Bimotrical geneticsChapman & Hall, London.

Morgan, J.P. 1988. Polycross designs with complete neighbor balance. Euphytica. 39:59-63. doi:10.1007/bf00025112.

Mosa, H.E. 2010. Estimation of combining ability of maize inbred lines using topcross mating design. J. Agricul. Res. Kafer El-Sheikh University. 36: 1-14.

Panhwar, S.A., M.J. Baloch, W.A. Jatoi, N.F. Veesar and M.S. Majeedano. 2008. Combining ability estimates from line x tester mating design in upland cotton. Pak. Acad. Sci. 45: 69-74.

Rashid, M., A.A. Cheema and M. Ashraf. 2007. Line x tester analysis in Basmati rice. Pak. J. Bot. 39: 2035-2042.

Schlegel, R.H.J. 2010. Dictionary of plant breeding. 2 ed. CRC Press,Taylor & Francis Group, Boca Raton.

Sharma, J.R. 2006. Statistical and biometrical techniques in plant breeding. 1 ed. New Age International. New Delhi.India

Shattuck, V.I., B. Christie and C. Corso. 1993. Principles of Griffing's combining ability analysis. Genetica. 90: 73-79.

Sleper, D.A. and J.M. Poehlman. 2006. Breeding field crops. 5th ed. Blackwell Publshing, Iowa, USA.

Wricke, G. and W.E. Weber. 1986. Quantitative genetics and selection in plant breeding.Walter de Gruyter & Co., Berlin.